Sampling signals with finite rate of innovation: the noisy case

Andrea Ridolfi [1], Irena Maravić [1], Julius Kusuma [2], Martin Vetterli [1,3]

École Polytechnique Fédérale de Lausanne - EPFL
School of Computer and Information Sciences

Technical Report 200285
December 23, 2002 (ID: IC/2002/85)

Abstract

In [1] a sampling theorem for a certain class of signals with finite rate of innovation (which includes for example stream of Diracs) has been developed. In essence, such non band-limited signals can be sampled at or above the rate of innovation. In the present paper, we consider the case of such signals when noise is present. Clearly, the finite rate of innovation property is lost, but if the signal-to-noise ratio (SNR) is sufficient, several methods are possible to reconstruct the signal while sampling well below the Nyquist rate. We thus explore the trade-offs between SNR, sampling rate, computational complexity and reconstruction quality. Applications of such methods can be found in acquisition and processing of signals in high bandwidth communications, like ultra wide band communication [2].

I. Introduction

Band-limited functions are just an example of signals that are specified by a fixed number of samples per unit of time, namely, if a signal $x(t)$ is band-limited to $[-\omega_m, \omega_m]$, then the famous sampling theorem by Shannon [3] states that $x(nT)$ with $T < \pi/\omega_m$, uniquely specifies the signal.

In [1], it is showed that certain signals with finite rate of innovation (a finite number of degrees of freedom per unit of time) could also be uniquely represented by uniform sampling with an appropriate kernel (*e.g.* the sinc kernel) and a rate just above the rate of innovation. Note that this sampling rate can be vastly inferior to the standard Nyquist rate, thus leading to very efficient, critical sampling schemes. To demonstrate the difference, consider a bi-level signal as used in CDMA. The Nyquist rate is given by the chip rate (the rate at which the signal changes levels) while the symbol rate corresponds to the much slower rate of chip rate divided by code length. This can be two or more order of magnitude slower (*e.g.* codes can be of length 256 or more). Of course, when there is noise, the deterministic theory developed so far is not applicable anymore. Indeed, noise is not a finite degrees of freedom signal. However, by taking more samples than necessary in the noiseless case, and applying appropriate estimation methods, it is possible to recover the underlying signal. The problem is closely related to parametric signal estimation as for example sinusoid retrieval in noise [4]. The interesting question is to understand the trade-offs between

1 Oversampling: how much above the critical sampling rate is it necessary to sample?
2 Algorithm: several methods are possible, from non-linear least squares to various subspace methods. Which one is best?
3 Computational complexity: as oversampling grows, how does the computational load evolve for the various methods?

1 Ecole Polytechnique Fédérale de Lausanne, LCAV - Laboratoire de Communications AudioVisuelles, 1015 Lausanne, Switzerland; {andrea.ridolfi,irena.maravic,martin.vetterli}@epfl.ch

2 Massachusetts Institute of Technology, LIDS - Laboratory for Information and Decision Systems, Cambridge, MA, USA; kusuma@mit.edu

3 University of California at Berkeley, EECS - Electrical Engineering and Computer Science Department, Berkeley, CA, USA.

4 Signal-to-noise ratio: for what ranges of noise can the signal still be recovered, and when do certain methods break down?

These questions are pursued below, in order to develop retrieval methods for finite-rate of innovation signals buried in noise. In particular, we focus on the retrieval of streams of weighted Diracs in noise. A direct application of these results are practical algorithms for sampling ultra wide-band signals at rates well below the “standard” Nyquist rate.

The outline of the paper is as follows.

Section II briefly reviews the basic idea of sampling signals with finite rate of innovation. We show the basic method in the noiseless case when the signal is a stream of weighted Diracs, namely the annihilating filter method.

Section III introduces the various methods that are possible when the Diracs are buried in noise, namely,

- subspace methods (state-space method, ESPRIT, MUSIC);
- multidimensional search methods (non linear least squares and optimization of Bernoulli-Gaussian models)

Section IV compares the methods for retrieval of Diracs in noise, giving a synthetic picture of when what method is usable, and at what cost.

Finally, Section V describes the application of this method to the problem of ultra wide band communication.

II. THE NOISELESS CASE

The sampling theory developed in [1] focuses on a particular class of signal with a finite rate of innovation, namely streams of weighted Diracs and piecewise polynomials. The core result can be summarized by the following theorem

Theorem II.1 [1] Sampling Theorem for signal with finite rate of innovation.
Let $x(t)$ be a τ periodic signal with D degrees of freedom (generally non band limited). Define the $[-D, D]$ low-pas approximation $y(t)$ of $x(t)$ as the result of its convolution with a filter that has zero Fourier coefficients except the ones with indexes $-D, \ldots, 0, \ldots, D$. Then, these $2D+1$ non-zero Fourier coefficients are a sufficient representation of the signal.
Such coefficients can be computed from $2D+1$ samples of the low-pass approximation uniformly taken at a rate $T = \tau/(2D+1)$.

The theorem also holds for discrete-time signal, where the non zero coefficients are computed from sub samples of the low pass approximation (see [1]). However, without loss of generality, in the following we refer to the periodic continuous time case.

In the particular case of a stream of K weighted Diracs,

$$x(t) = \sum_{k=1}^{K} \alpha_k \delta(t - t_k) \tag{1}$$

positions $\{t_k\}$ and weights $\{\alpha_k\}$ are sequentially computed. Due to their linear character, the weights are obtained in closed form once the positions of the Diracs are known, while the positions are retrieved by means of the annihilating filter method.

The **annihilating filter method** consists in finding the coefficients of the $(D+1)$-th order filter that annihilates the data, *i.e.* the non-zero Fourier values of the sampled low-pass approximation

$$y_n = \int_0^\tau h(t - nT)\, x(t)\, dt, \quad n \in \mathbb{N}$$

where $h(t)$ is the low-pass filter, *i.e.* the sampling kernel. The annihilating filter exists and is unique, up to a proportional term. Its zeros bear complete information about the positions of the Diracs and can be retrieved by a root-finding operation.

In practice, the annihilating filter can be seen as a subspace method since its coefficients are computed by means of an eigendecomposition of the Toeplitz matrix associated to the $2D+1$ non-zero Fourier values.

III. The noisy case

We now consider a stream of weighted Diracs $x(t)$ affected by an additive Gaussian white noise $\epsilon(t)$ and we suppose to have samples of its low pass approximation

$$y_n = \int_0^\tau h(t - nT)(x(t) + \epsilon(t))\, dt, \quad n \in \mathbb{N}$$

Theorem II.1 does not apply for the simple reason that the noise is not a signal with finite rate of innovation. However, retrieving a stream of weighted Diracs can be still performed sequentially, in the sense that weights can be computed in closed form once the positions are known. Thus, the problem lies in finding noise-robust methods for retrieving the positions of the Diracs.

Parametric line spectra estimation theory [4], harmonic retrieval [5] or stochastic modeling approaches [6] provide several noise-robust methods. In particular we consider the ones described in the following.

A. Subspace methods

Subspace methods derive position estimates by exploiting the properties of the eigendecomposition of matrices related to the Fourier data and, in particular, the subspaces associated with those matrix. Their most attractive feature is that they are based on relatively simple matrix manipulations.

Due to the root-finding step, the annihilating filter method that we have introduced in Section II may have poor performances in presence of noise. Other more noise-robust subspace methods are available: in [5] a "state-space" method is proposed which can be adapted to our problem, while line spectra estimation theory [4] provides well developed methods such as ESPRIT and MUSIC.

In spite of the common subspace theme between annihilating filter and state-space methods, on one side, and ESPRIT and MUSIC methods, on the other, they are substantially different. Indeed, the latter two methods deal with second order quantities of the signal (spectral analysis): the model of the signal is based on the covariance matrix of the data ("covariance matrix model" [4]). In such a model, the noise is intrinsically taken into account through its second order properties. The eigenstructure of the covariance matrix of the Fourier data contains complete information on the positions, thus is the object of the eigendecomposition. More precisely:

The **state-space** method retrieves the position by exploiting a shift invariance property of particular subspaces of the Fourier data matrix;

ESPRIT (Estimation of Signal Parameters by Rotational Invariance Technique) retrieves the position by exploiting a shift invariance property of particular subspaces of the covariance matrix;

MUSIC (MUltiple SIgnal classification) retrieves the position by exploiting the fact that the covariance matrix of the Fourier data corresponding to the noiseless stream of Diracs and the subspace associated to the noise are orthogonal.

B. Multidimensional search methods

Non linear least squares and Bernoulli-Gaussian approaches are based on the optimization of a criterion with respect to the positions of the Diracs. More precisely:

The **non linear least squares** method estimates positions as the minimizers of the square error

between the Fourier data and the signal model in the Fourier domain. When the noise is supposed white and Gaussian, such a method can be interpreted as a maximum likelihood estimation.
The **Bernoulli-Gaussian** approach consists in modeling the signal as a Bernoulli-Gaussian process and then in estimating the positions as the maximizers of the related likelihood function [6].

Within the maximum likelihood framework, numerical optimization can be globally performed by means of stochastic algorithms, such as simulated annealing, or locally performed by semi-stochastic or deterministic algorithm. Note that in the case of a small number of Diracs, simulated annealing with a finite annealing schedule [7] may be an attractive approach. However, in the general case the computational burden of such numerical optimization methods is so high that in practice they are unusable.

IV. Numerical results and comparison

In order to illustrate the numerical performances of the methods we have presented, we have considered discrete time streams of 2 weighted Diracs in a length 512 signal. The positions are randomly chosen according to a uniform distribution over $[1, 512]$, while the weights are sampled from i.i.d. zero mean Gaussian random variables with unit variance. Therefore, the signal power is equal to 2. The noise is supposed white and Gaussian. Each method is tested for different values of the noise variance and different values of the low-pass filter bandwidth, *i.e.* for different number of samples. The mean square estimation error of each method is computed for one thousands signals (ten stream of weighted Diracs, each one added to one hundred realizations of the noise) and then averaged.

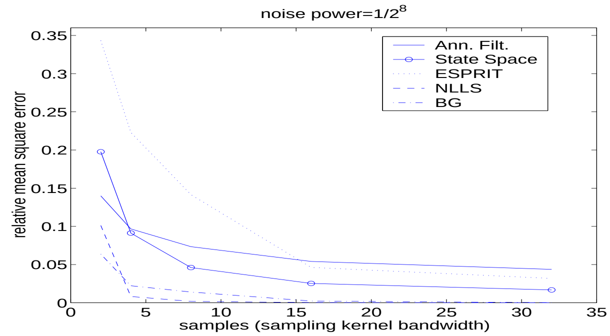
As stated in [4], ESPRIT and MUSIC have similar statistical accuracy and in our simulations they indeed provide very similar results. Therefore, only the ones obtained with ESPRIT are shown.

The results of the non linear least squares and Bernoulli-Gaussian model approaches have been obtained with an exhaustive optimization search: they are to be considered as a benchmark for the comparison with the other numerical methods.

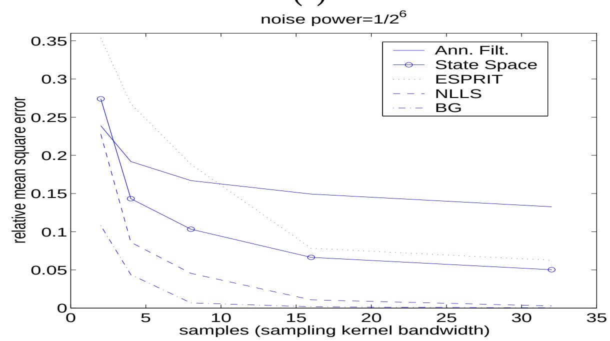
Figure 1 shows relative mean square errors of position estimates versus the number of samples, *i.e.* the sampling kernel bandwidth, for two different values of the noise variance: 2^{-8} and 2^{-6}. Note that such values correspond to a signal-to-noise ratio of 0db and -13.9db, respectively. The performance of the annihilating filter is strongly noise dependent and it slightly improve when the number of samples increases. From our simulation of ESPRIT, it is evident that, when the number of available samples is small, the results are extremely poor (worse than the annihilating filter ones!). This is due to the fact that ESPRIT requires the estimation of the covariance matrix which critically depends on the number of samples. The performance of ESPRIT improves when the number of samples increases, outperforming the annihilating filter (see Figure 1 (b)). The state-space method has a good overall performance, better than the two other subspace methods, except for the critical number of samples $2K+1$, where it is outperformed by the annihilating filter. However, as shown in Figure 2, the price of such a better performance is a higher complexity, compared to the previous two methods. These subspace methods present a “break” point in the error curve. The error of annihilating filter and the state-space methods “breaks” at $8K+1$ samples (*i.e.* 2^2 times the critical number) when the noise power is equal to 2^{-8} (Figure 1 (a)), and at $16K+1$ samples (*i.e.* 2^3 times the critical number) when the noise power is equal to 2^{-6} (Figure 1 (b)),

As expected, the non linear least squares method and the optimization of a Bernoulli-Gaussian model yield the best results in the presence of noise and clear “break” points in the error curves. Among the two approaches, the Bernoulli-Gaussian is even more noise robust, as shown in Figure 1 (b). However, as already mentioned, the price to pay in terms of complexity is generally too high for communication applications.

Figure 2 depicts the complexity versus the number of samples, *i.e.* the sampling kernel bandwidth, for the subspace methods



(a)



(b)

Fig. 1. Relative mean square error versus the number of samples (sampling kernel bandwidth) for all the explored methods. The top and bottom figures depict the results for a noise variance respectively equal to 2^{-8} and 2^{-6}

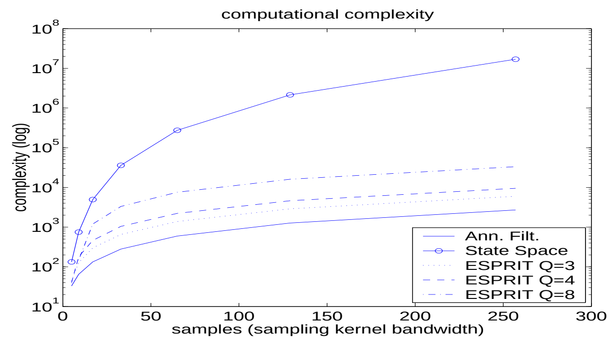


Fig. 2. Complexity versus number of samples (sampling kernel bandwidth) of the subspace methods

V. Sampling ultra wide band signals

We now turn our attention to the sampling of Ultra-Wideband (UWB) signals. UWB signals are designed to have very high bandwidth expansion, where the transmitted signal's bandwidth is many times the symbol rate. This in turn poses a challenge for the receiver, as the Nyquist rate required to sample the signal is exceedingly high. We note that the true innovation rate of the signal is dictated by the symbol rate. The framework developed in this work allows one to reliably sample close to the innovation rate or the symbol rate. This was first presented in [2], where it was also demonstrated that noise robustness increases as a function of oversampling beyond the critical rate.

The main idea is as follows: we interpret UWB signals as a train of Diracs carrying the information at the symbol rate. This train of Diracs is then convolved with the pulse shape, and in turn subjected to the noisy effects of a linear time-invariant channel, and receive filter. We can jointly consider these effects and refer to them as the *compound channel*, as shown in Figure 3.

Therefore, through equalization, the framework for this problem derives from the framework for

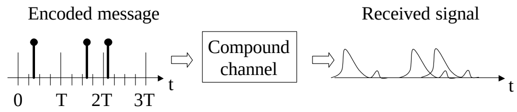


Fig. 3. Illustration of the compound channel

sampling of signals with a finite rate of innovation with a linear time-invariant kernel [1].

With the results of Section IV, we see that subspace methods with sufficient oversampling can be used to detect UWB signals using sampling well below the Nyquist rate.

VI. Concluding remarks

We have explored the retrieval of noisy weighted Diracs in the light of the recent work on sampling noiseless signals with finite rate of innovation [1]. The performances of subspace and multidimensional search methods have been investigated. Numerical results show that sampling signal with finite rate of innovation at a rate below the standard Nyquist rate, but above the critical rate, is still possible in the noisy case.

Therefore, the framework in [1] has been shown to be applicable in the noisy case as well, and can be used for communication problems like UWB communications. With increasing interest in bandwidth-expanding short range communication systems, this work gives a promising framework with which communication engineers can trade-off system performance with complexity, using sampling well below the Nyquist rate.

References

[1] M. Vetterli, P. Marziliano, and T. Blu, “Sampling signals with finite rate of innovation,” *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1417–1428, June 2002.
[2] J. Kusuma, A. Ridolfi, and M. Vetterli, “Sampling of communications systems with bandwidth expansion,” in *Proc. IEEE ICC 2002*.
[3] C. E. Shannon, “A mathematical theory of communication,” *Bell System Tech. J.*, vol. 27, pp. 379–423,623–656, 1948.
[4] P. Stoica and R. Moses, *Introduction to Spectral Analysis*. Prentice Hall, 1997.
[5] S. Y. Kung, K. S. Arun, and D. V. Bhaskar Rao, “State-space and singular-value decomposition-based approximation methods for the harmonic retrieval problem,” *J. Opt. Soc. Amer.*, vol. 73, no. 12, pp. 1799–1811, December 1983.
[6] J. J. Kormylo and J. M. Mendel, “Maximum-likelihood detection and estimation of Bernoulli-Gaussian processes,” *IEEE Trans. Inform. Theory*, vol. 28, pp. 482–488, 1982.
[7] O. Catoni, “Solving scheduling problems by simulated annealing,” *SIAM J. Control Optim.*, vol. 36, no. 5, pp. 1539–1575, 1998.